Laboration 3 - Icke-linjär optimering

Laborationens första del behandlar lösning av begränsade icke-linjära problem och belyser även konsekvenserna av icke-konvexitet. Den andra delen syftar till att ge insikt i hur metoder för obegränsad optimering (Brantaste lutningsmetoden och Newtons metod) uppträder för konvexa och icke-konvexa problem, samt att ge insikt i hur straffunktionsmetoder fungerar. Laborationen utnyttjar MATLAB.

I ett terminalfönster ger kommandona

```
module add prog/matlab/
matlab &
```

Inne i MATLAB-fönstret ges kommandona

```
addpath /courses/TAOP07/matlab/
startaILP
```

Begränsad optimering

1. Givet följande problem.

$$\begin{array}{llcl} \min & f(x) & = & 2(x_1-3)^2+(x_2-2)^2-3x_2-x_1x_2 \\ \text{då} & x_1^2 - x_2 & \leq & 0 \qquad (1) \\ & -x_1 + x_2 & \leq & 2 \qquad (2) \end{array}$$

Problemet finns i MATLAB-programmet som **Constrained ILP 1**. En grafisk illustration av problemet finns i figuren nedan.

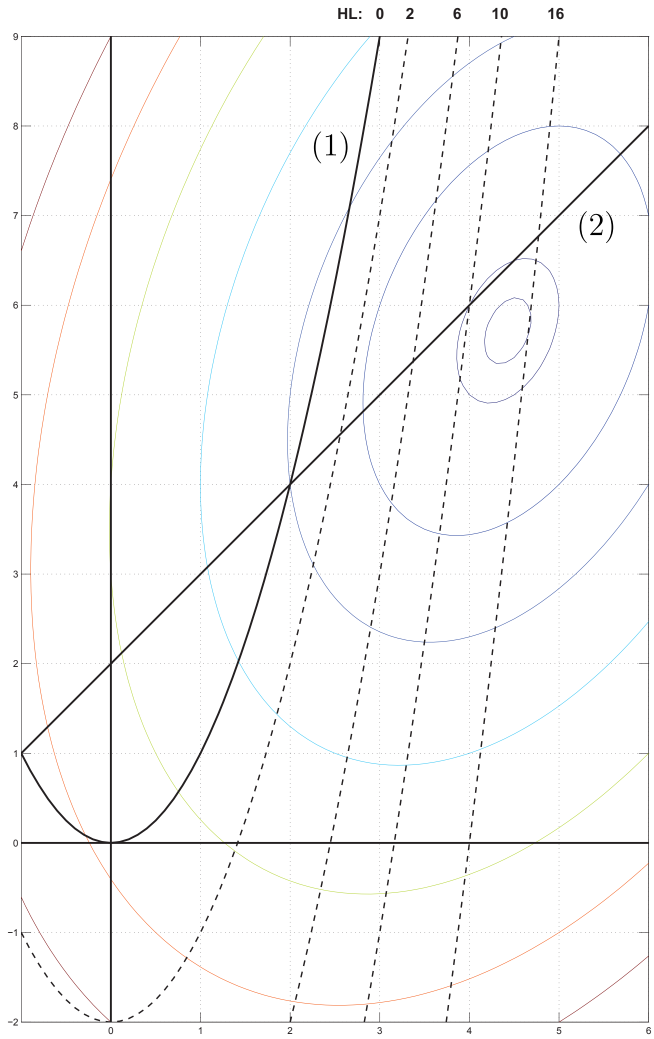
Lös problemet. Ge optimallösningen och markera den i figuren.

. . . .

Ge multiplikatorernas (dualvariablernas) värden.

. . . .

Illustrera *grafiskt* i figuren att den funna punkten uppfyller KKT-villkoren.



(**Ska förberedas**.) Undersök om problemet är konvext, för att avgöra om lösningen är ett *globalt* minimum.

. . . .

. . . .

. . . .

. . . .

. . . .

. . . .

Ändra det *första* bivillkoret till $x_1^2 - x_2 \leq 2$. Hur stor förändring av det optimala målfunktionsvärdet kan förväntas, utifrån värdet på bivillkorets multiplikator?

. . . .

Lös om problemet och markera den nya lösningen i figuren. Vilket optimalt målfunktionsvärde fås? Förklara skillnaden mellan detta värde och det som indikeras av multiplikatorvärdet ovan. Vad är nya värdet på multiplikatorn för det första bivillkoret?

. . . .

. . . .

. . . .

. . . .

Lös även problemet för högerleden 6, 10 och 16. Markera optimallösningarna i figuren. Hur förändras värdet på bivillkorets multiplikator?

. . . .

. . . .

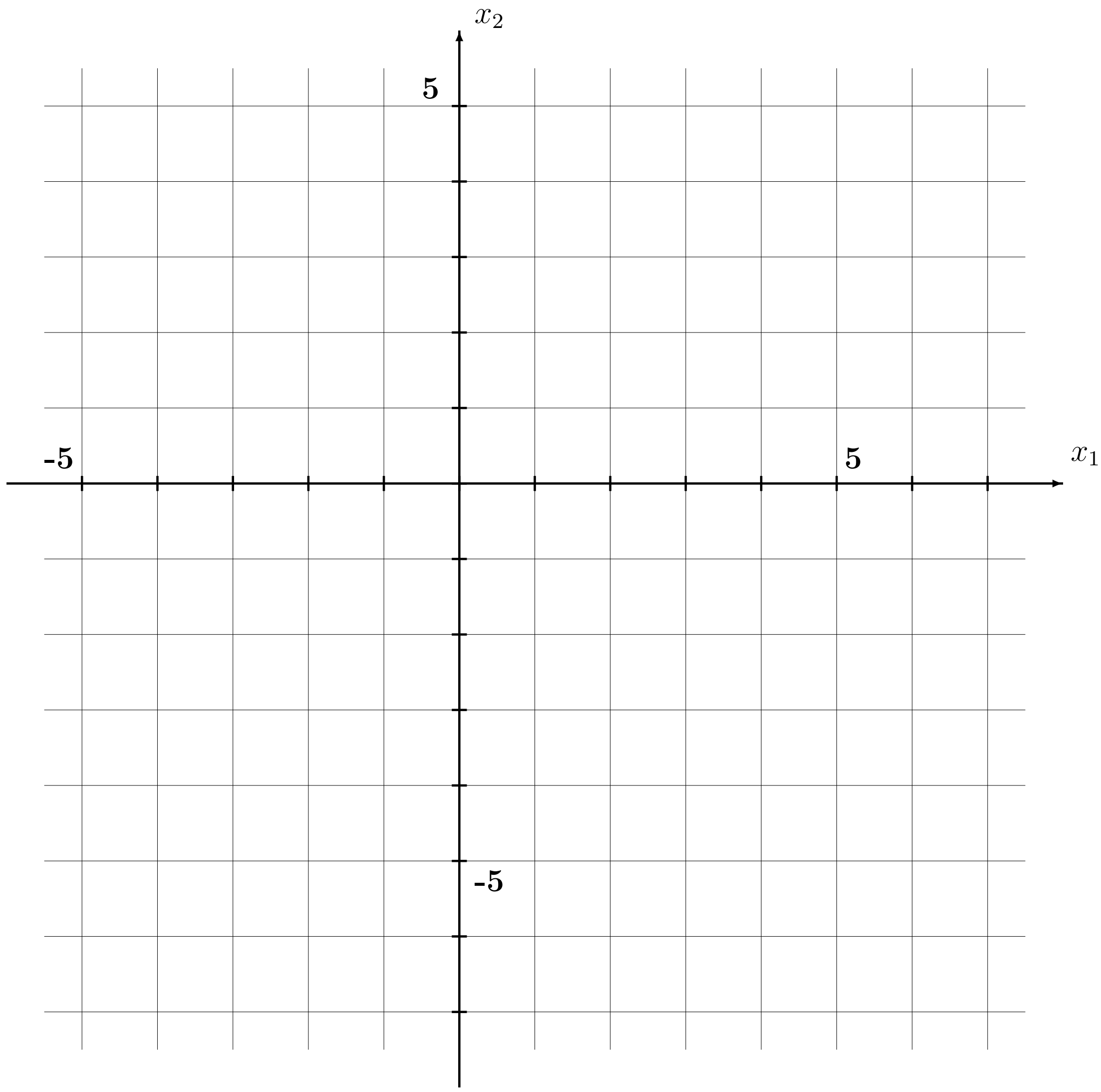
. . . .

2. Givet problemet

$$
\begin{array}{llcccl}
\min & f(x_1, x_2) = x_1 & & & & \\
\text{då} & (x_1 - 1)^2 & + & (x_2 + 2)^2 & \leq & 16 \\
 & x_1^2 & + & x_2^2 & \geq & 13,
\end{array}
$$

vilket är **Constrained ILP 2**.

(**Ska förberedas**.) Rita det tillåtna området i figuren nedan. Vilka punkter uppfyller KKT-villkoren? Motivera *grafiskt*. Vilka punkter är *lokala* minima? Vilken punkt är *globalt* minimum? Motivera noga!


x_2
5
-5
5
x_1
-5

....

....

....

....

....

Lös problemet. Starta från punkterna (0,4), (-4,4) och (3.6,0), samt från ett par andra punkter. (Även otillåtna punkter går bra att starta ifrån.) Vilka lösningar fås?

. . . .

Obegränsad optimering

De funktioner som skall studeras finns redan inskrivna, och det som återstår är att välja rätt funktion och metod. Det finns två Newton-metoder att välja på: en enkel Newtons metod som alltid tar ett ett-steg i Newton-riktningen, och Newtons *modifierade* metod, som gör en linjesökning i Newton-riktningen. I **alla** uppgifter ska funktionen **minimeras**.

3. Studera **funktion 1**

$$f(x_1, x_2) = 2(x_1 + 1)^2 + 8(x_2 + 3)^2 + 5x_1 + x_2$$

Lös med Brantaste lutningsmetoden (BL) och Newtons (modifierade) metod. Starta i punkterna $(10, 10)$ och $(-5, -5)$, samt i någon egen startpunkt. Mot vilken punkt konvergerar metoderna?

. . . .

Är den uppnådda punkten ett globalt minimum? Varför (inte)?

. . . .

För vilken problemtyp konvergerar Newtons metod alltid på *en* iteration (för varje startpunkt)?

. . . .

4. Studera **funktion 2** (Rosenbrocks funktion)

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

Lös med BL och Newtons (modifierade) metod. Starta i punkten $(-1.5, -1)$. Mot vilken punkt konvergerar metoderna (om ett stort antal iterationer utförs)?

. . . .

Är funktionen konvex? . . . .

Konvergerar metoderna i detta fall mot ett globalt optimum? ....

Studera gärna metodernas uppträdande för några andra startpunkter.

5. Studera **funktion 3**

$$
\begin{aligned}
f(x_1,x_2) = \quad & -6e^{-\frac{(x_1+1)^2+(x_2+1)^2}{10}} + 4e^{-\frac{x_1^2+x_2^2}{100}} - 5e^{-\frac{(x_1-6)^2+(x_2+2)^2}{20}} - \\
& -5e^{-\frac{(x_1+3)^2+(x_2-5)^2}{30}} - 4e^{-\frac{(x_1+3)^2+(x_2-5)^2}{300}}
\end{aligned}
$$

Lös med BL och starta i punkterna (-15,10), (10,-10), (5,4), (15,0) och (-1,-10). Mot vilka punkter konvergerar metoden?

....

För att en Newton-riktning säkert ska ge ett minskande målfunktionsvärde måste funktionen vara strikt konvex (Hessianen positivt definit) i en omgivning till den punkt där vi beräknar riktningen. Är detta inte fallet så kan riktningen istället komma att ge ett ökande målfunktionsvärde, vilket medför att den optimala steglängden i linjesökningen blir noll (eftersom vi där söker minimum med avseende på en icke-negativ steglängd). Ett sätt att alltid få en avtagande-riktning i en Newton-metod, oberoende av funktionens konvexitets-egenskaper, är att se till så att matrisen som används vid sökriktningsbestämningen *alltid* är positivt definit. Det kan göras genom att man om Hessianen $\nabla^2 f(x^k)$ inte är positivt definit, adderar matrisen νI till $\nabla^2 f(x^k)$, där I är en enhetsmatris av lämplig storlek och ν är strikt större än beloppet av det minsta egenvärdet till $\nabla^2 f(x^k)$. Detta ger en sökriktning

$$
d^k = -\left(\nabla^2 f(x^k) + \nu I\right)^{-1} \nabla f(x^k),
$$

vilken alltid resulterar i ett minskande målfunktionsvärde. Denna modifierade metod kallas Newton–Marquardt. [Notera att då ν växer så kommer termen νI alltmer dominera över $\nabla^2 f(x^k)$, vilket gör att d^k blir alltmer parallell med brantaste lutningsriktningen (dvs $-\nabla f(x^k)$).]

(**Ska förberedas**.) Visa att om λ är ett egenvärde till $\nabla^2 f(x^k)$ så är $\gamma = \lambda + \nu$ ett egenvärde till $\nabla^2 f(x^k) + \nu I$. [Ledning: Använd sekularekvationen.]

....

[Om $\nabla^2 f(x^k)$ har något negativt egenvärde och ν är strikt större än beloppet av det *mest* negativa egenvärdet, så fås då att $\nabla^2 f(x^k) + \nu I$ är positivt definit och att $\nabla f(x^k)^{\mathrm{T}} d^k = -[(\nabla^2 f(x^k) + \nu I) d^k]^{\mathrm{T}} d^k = -(d^k)^{\mathrm{T}} (\nabla^2 f(x^k) + \nu I) d^k < 0$, dvs att d^k då är en avtagande-riktning.]

Lös med Newtons (modifierade) metod och Newton–Marquardt från startpunkten (-15,15).

Här fungerar inte en Newton-metod; varför inte? ....

....

6. Studera **funktion 4**

$$f(x_1, x_2) = x_1^3 + 2x_2^3 - 27x_1 - 6x_2$$

(**Ska förberedas**.) Beräkna Hessianen till funktionen.

Starta i punkten (0,0) och lös med BL, Newtons (modifierade) metod och Newton–Marquardt.

Varför fungerar inte en Newton-metod? ....

....

Starta i några valfria punkter och försök hitta så många stationära punkter som möjligt. Hur många hittar ni?

....

Vilken typ är dessa? ....

7. Givet problemet i uppgift 2.

(**Ska förberedas**.) Använd en *yttre* straffunktion för att omformulera till ett obegränsat optimeringsproblem.

....

....

Studera denna funktion under **PenaltyFunc ILP2**. Lös med BL. Använd straffparameter $\mu = 0.005$ och starta i (2,0).

Vad blir optimallösningen (med tre decimaler)? ....

Lös med straffparameter $\mu = 0.01$ och med start från den lösning som just beräknats.

Vad blir optimallösningen nu? ....

Upprepa med straffparametrar $\mu = 0.1$ och $\mu = 1$. Vilka optimallösningar fås?

....

8. Studera i mån av tid övriga funktioner som finns i menyn.

   Funktion 5: En "dalgång". Starta tex i (0.8, 0.3) med Newtons (modifierade) metod.

   Funktion 7: Många "gropar".

   Funktion 8: Med "kullar och sänkor".